Donald L. Phillips · Jillian W. Gregg

Uncertainty in source partitioning using stable isotopes

Abstract Stable isotope analyses are often used to quantify the contribution of multiple sources to a mixture, such as proportions of food sources in an animal's diet, or C_3 and C_4 plant inputs to soil organic carbon. Linear mixing models can be used to partition two sources with a single isotopic signature (e.g., $\delta^{13}C$) or three sources with a second isotopic signature (e.g., $\delta^{15}N$). Although variability of source and mixture signatures is often reported, confidence interval calculations for source proportions typically use only the mixture variability. We provide examples showing that omission of source variability can lead to underestimation of the variability of source proportion estimates. For both two- and three-source mixing models, we present formulas for calculating variances, standard errors (SE), and confidence intervals for source proportion estimates that account for the observed variability in the isotopic signatures for the sources as well as the mixture. We then performed sensitivity analyses to assess the relative importance of: (1) the isotopic signature difference between the sources, (2) isotopic signature standard deviations (SD) in the source and mixture populations, (3) sample size, (4) analytical SD, and (5) the evenness of the source proportions, for determining the variability (SE) of source proportion estimates. The proportion SEs varied inversely with the signature difference between sources, so doubling the source difference from 2‰ to 4‰ reduced the SEs by half. Source and mixture signature SDs had a substantial linear effect on source proportion SEs. However, the population variability of the sources and the mixture are fixed and the sampling error component can be changed only by increasing sample size. Source proportion SEs varied inversely with the square root of sample size, so an increase from 1 to 4 samples per population cut the SE in half. Analytical SD had little effect over the range examined since it was generally substantially smaller than the population SDs. Proportion SEs were minimized when sources were evenly divided, but increased only slightly as the proportions varied. The variance formulas provided will enable quantification of the precision of source proportion estimates. Graphs are provided to allow rapid assessment of possible combinations of source differences and source and mixture population SDs that will allow source proportion estimates with desired precision. In addition, an Excel spreadsheet to perform the calculations for the source proportions and their variances, SEs, and 95% confidence intervals for the two-source and three-source mixing models can be accessed at http://www.epa.gov/wed/pages/models.htm.

Introduction

The use of stable isotope analyses in ecology has increased dramatically in the last several decades (Griffiths 1998). One use of stable isotopes is to determine the proportional contributions of several sources to a mixture. Examples of source proportion calculations include determination of: various food sources in an animal's diet (Szepanski et al. 1999); soil carbon derived from current and previous vegetation (Vitorello et al. 1989); water sources used by plants (Dawson 1993); percent of N derived from N-fixation (Hogberg 1997); and sources of CO_2 efflux from forest floor respiration (Lin et al. 1999). Linear mixing models are used to estimate proportions for two sources using isotopic signatures for a single element (e.g., $\delta^{13}C$; Balesdent and Mariotti 1996), or for three sources using isotopic signatures for two elements (e.g., $\delta^{13}C$ and $\delta^{15}N$; Phillips 2001).

To determine the relative proportion of each source to the mixture, isotopic signatures are typically determined for multiple samples of each source and the mixture. Mean signature values are calculated for each of the sources, and the fractional contributions of each source to

D.L. Phillips (✉) · J.W. Gregg
U.S. Environmental Protection Agency,
National Health and Environmental Effects Research Laboratory,
200 SW 35th Street, Corvallis, OR 97333, USA
e-mail: don@mail.cor.epa.gov

the mixture are calculated using these point estimates. Calculations may also use a mean isotopic signature for the mixture, or perform a separate calculation for each sample of the mixture and report confidence intervals for fractional contributions of the sources based on this replication. However, the confidence intervals are based solely on the variability of the isotopic signatures of the mixture, and do not include the variability in the isotopic composition of the individual sources. In general, little effort has been made to enumerate various factors that contribute to uncertainty and to determine how those errors propagate into uncertainty about the estimates of source proportions.

Veldkamp and Weitz (1994) provide a good example of an uncertainty analysis for a two-source mixing model using stable isotopes. They performed Monte Carlo simulations to determine how variability in various parameter estimates affected the uncertainty of estimates of the proportions of soil organic matter derived from previous and current vegetation. While they described a valuable example of error analysis for a particular case study, it would be useful to have an equation to provide an analytical solution for the uncertainty of source proportions, rather than having to estimate this through hundreds or thousands of Monte Carlo simulations for each different case. The purpose of this paper is: (1) to provide formulas for approximate variances, SEs, and confidence intervals for source proportions that account for variability in the isotopic signatures of both the sources and the mixture; (2) to determine the sensitivity of the source proportion variances to a variety of contributing factors over a range of conditions; and (3) to use the results of our sensitivity analyses to reach some conclusions about where efforts at reducing uncertainty of source proportions should be focused. Specifically, we show how our approach is useful for determining the required sample sizes and isotopic signature difference between sources to obtain source proportion estimates with a specified level of precision.

Methods

Calculating variances and confidence intervals

Single isotope, two-source mixing model

Figure 1a shows an example of the analytical situation for the single isotope, two-source mixing problem. Random samples are drawn from the populations of each of the sources (A and B) and of the mixture (M). For our purposes, we assume that either there is no fractionation in incorporating a source into the mixture, or the isotopic signatures for the sources have already been adjusted to account for this fractionation, as is frequently done in dietary analyses (e.g., Szepanski et al. 1999). A two-endmember linear mixing model can be formulated from the following mass balance equations:

$$\bar{\delta}_M = f_A \bar{\delta}_A + f_B \bar{\delta}_B$$
$$1 = f_A + f_B \tag{1}$$

and the mean proportion of source A in the mixture can be calculated (Balesdent and Mariotti 1996) as:

$$f_A = \frac{\bar{\delta}_M - \bar{\delta}_B}{\bar{\delta}_A - \bar{\delta}_B} \tag{2}$$

where $\bar{\delta}_M$, $\bar{\delta}_A$, and $\bar{\delta}_B$ represent the mean isotopic signatures (e.g., δ^{13}C) for the mixture M and sources A and B, respectively, and f_A and f_B are the proportions of A and B in M. If we make the reasonable assumption that $\bar{\delta}_M$, $\bar{\delta}_A$, and $\bar{\delta}_B$ are independently measured, then a first-order Taylor series approximation of the variance of f_A evaluated at ($\bar{\delta}_M$, $\bar{\delta}_A$, and $\bar{\delta}_B$) can be calculated using partial derivatives (Taylor 1982) as:

$$\sigma_{f_A}^2 = \left(\frac{\delta f_A}{\partial \bar{\delta}_M}\right)^2 \sigma_{\bar{\delta}_M}^2 + \left(\frac{\delta f_A}{\partial \bar{\delta}_A}\right)^2 \sigma_{\bar{\delta}_A}^2 + \left(\frac{\delta f_A}{\partial \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_B}^2 \tag{3}$$

which reduces to (see Appendix):

$$\sigma_{f_A}^2 = \frac{1}{(\bar{\delta}_A - \bar{\delta}_B)^2}\left[\sigma_{\bar{\delta}_M}^2 + f_A^2 \sigma_{\bar{\delta}_A}^2 + (1 - f_A)^2 \sigma_{\bar{\delta}_B}^2\right] \tag{4}$$

where $\sigma_{\bar{\delta}_M}^2$, $\sigma_{\bar{\delta}_A}^2$, and $\sigma_{\bar{\delta}_B}^2$ represent variances of the mean isotopic signatures for the mixture M, and sources A and B, respectively (i.e., the square of the SEs). An approximate variance for $f_B = 1 - f_A$ can also be determined by switching the A and B subscripts in Eq. 4. Approximate 95% confidence intervals for f_A can be constructed as $f_A \pm t_{0.05,\gamma}\,\sigma_{fA}$ (and similarly for f_B), where σ_{fA} is the SE of the proportion estimate (square root of the variance from Eq. 4), and $t_{0.05,\gamma}$ represents the two-tailed Student's t for $\alpha=0.05$ and γ degrees of freedom. γ represents the Satterthwaite (1946) approximation for the degrees of freedom associated with σ_{fA}^2, which can be calculated as:

$$\gamma = \frac{\left(\sum_i c_i V_i\right)^2}{\sum_i \frac{(c_i V_i)^2}{d_i}} \tag{5}$$

where V_i, c_i, and d_i are the individual variance terms, their coefficients, and their associated degrees of freedom (n_i-1) on the right side of Eq. 4. Note that the mean isotopic signature and its SE for a source or mixture may also be derived as the y-intercept (and its SE) from linear regressions of n_i points in "Keeling plots" (δ vs. 1/concentration; Keeling 1958). In this case, since the linear regression is a two-parameter model, the degrees of freedom associated with this variance term in Eq. 5 should be adjusted to n_i-2. The proportions, their SEs, and their confidence intervals can also be expressed as percentages by multiplying by 100%.

Dual isotope, three-source mixing model

An example of the analytical situation for random samples of three sources (A, B, and C) and the mixture is shown in Fig. 1b. A dual isotope, three-endmember linear mixing model can be formulated from the following mass balance equations:

$$\bar{\delta}_M = f_A \bar{\delta}_A + f_B \bar{\delta}_B + f_C \bar{\delta}_C$$
$$\bar{\lambda}_M = f_A \bar{\lambda}_A + f_B \bar{\lambda}_B + f_C \bar{\lambda}_C \tag{6}$$
$$1 = f_A + f_B + f_C$$

and the mean proportion of source A in the mixture can be calculated (Phillips 2001) as:

$$\sigma_{f_A}^2 = \left(\frac{\partial f_A}{\partial \bar{\delta}_M}\right)^2 \sigma_{\bar{\delta}_M}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_A}\right)^2 \sigma_{\bar{\delta}_A}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_B}^2 \tag{7}$$

where $\bar{\delta}$ and $\bar{\lambda}$ represent mean isotopic signatures for two elements (e.g., δ^{13}C and δ^{15}N), and the subscripts refer to the sources A, B, and C, and the mixture M. If we again assume the independence of the isotopic signature measurements of the three sources and the mixture, then a first order Taylor series approximation for the variance of f_A evaluated at the mean is:

$$\sigma_{f_A}^2 = \left(\frac{\partial f_A}{\partial \bar{\delta}_M}\right)^2 \sigma_{\bar{\delta}_M}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_A}\right)^2 \sigma_{\bar{\delta}_A}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_B}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_C}\right)^2 \sigma_{\bar{\delta}_C}^2$$
$$+ \left(\frac{\partial f_A}{\partial \bar{\delta}_M}\right)^2 \sigma_{\bar{\lambda}_M}^2 + \left(\frac{\partial f_A}{\partial \bar{\lambda}_A}\right)^2 \sigma_{\bar{\lambda}_A}^2 + \left(\frac{\partial f_A}{\partial \bar{\lambda}_B}\right)^2 \sigma_{\bar{\lambda}_B}^2 + \left(\frac{\partial f_A}{\partial \bar{\lambda}_C}\right)^2 \sigma_{\bar{\lambda}_C}^2 \tag{8}$$

Fig. 1 a Partitioning of sugar cane field soil C into contributions from sugar cane and soil from the previous forest vegetation using δ^{13}C (Vitorello et al. 1989). δ^{13}C in per mil (‰) units is the deviation of the isotopic ratio of the sample from that for a PeeDee Belemnite standard; δ^{13}C = $(R_{\text{sample}}/R_{\text{standard}}-1)\times1000$, where $R = {}^{13}$C$/{}^{12}$C. **b** Partitioning of dietary contributions of moose, caribou, and salmon for interior Alaska wolves using δ^{13}C and δ^{15}N (Szepanski et al. 1999). δ^{15}N in per mil (‰) units is the deviation of the isotopic ratio of the sample from that for an atmospheric nitrogen standard; δ^{15}N = $(R_{\text{sample}}/R_{\text{standard}}-1)\times1000$, where $R = {}^{15}$N$/{}^{14}$N. The *numbers in parentheses* are the proportions as calculated from the mean isotopic signatures for the sources and the mixture using Eq. 4 for **a** and Eq. 9 for **b**. *Error bars* show ±1 SE (*inner bars*) and 95% confidence intervals ($\pm t_{0.05,df}$ SE; *outer bars*) for isotopic signatures
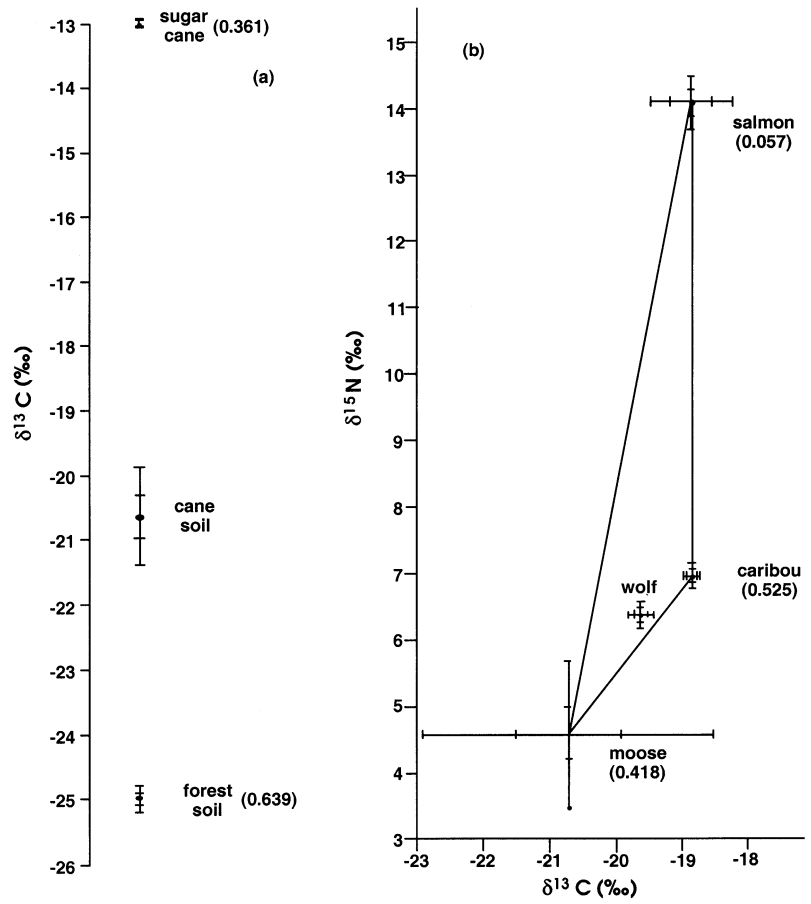
Letting N and D represent the numerator and denominator of Eq. 7, respectively, Eq. 8 reduces to:

$$\sigma_{f_A}^2 = \frac{1}{D^4}\left\{\left[D(\bar{\lambda}_M - \bar{\lambda}_C) - N(\bar{\lambda}_A - \bar{\lambda}_C)\right]^2 \sigma_{\bar{\delta}_B}^2 + \left[N(\bar{\lambda}_B - \bar{\lambda}_C)\right]^2 \sigma_{\bar{\delta}_A}^2 \right.$$
$$+ \left[D(\bar{\lambda}_B - \bar{\lambda}M) - N(\bar{\lambda}_B - \bar{\lambda}_A)\right]^2 \sigma_{\bar{\delta}_C}^2 + \left[D(\bar{\lambda}_C - \bar{\lambda}_B)\right]^2 \sigma_{\bar{\delta}_M}^2$$
$$+ \left[D(\bar{\delta}_C - \bar{\delta}_M) - N(\bar{\delta}_C - \bar{\delta}_A)\right]^2 \sigma_{\bar{\lambda}_C}^2 + \left[N(\bar{\delta}_C - \bar{\delta}_B)\right]^2 \sigma_{\bar{\lambda}_A}^2$$
$$\left. + \left[D(\bar{\delta}_M - \bar{\delta}_B) - N(\bar{\delta}_A - \bar{\delta}_B)\right]^2 \sigma_{\bar{\lambda}_C}^2 + \left[D(\bar{\delta}_B - \bar{\delta}_C)\right]^2 \sigma_{\bar{\lambda}_M}^2\right\} \quad (9)$$

While the measurements of the isotopic signatures for samples from the mixture population and three source populations should be independent of each other, within each population there may be correlation of the two isotopic signatures. For example, the δ^{18}O and δ^2H signatures in individual plant and soil water compartments are often highly correlated since both H and O are found in the water molecule. δ^{18}O and δ^2H signatures in precipitation tend to be linearly related and fall along a meteoric water line (Dawson 1998). If the two isotopic signatures are correlated within the mixture and source populations, covariance terms should be added to Eqs. 8, 9 to account for this. Eqs. 10, 11, 12 below show the calculations for proportion variances corrected for such correlation ($\sigma_{f_a}^{2'}$):

$$\sigma_{fA}^{2'} = \sigma_{f_A}^2 + 2\frac{\partial f_A}{\partial \bar{\delta}_M}\frac{\partial f_A}{\partial \bar{\lambda}_M}\sigma_{\bar{\delta}_M,\bar{\lambda}_M} + 2\frac{\partial f_A}{\partial \bar{\delta}_A}\frac{\partial f_A}{\partial \bar{\lambda}_A}\sigma_{\bar{\delta}_A,\bar{\lambda}_A}$$
$$+ 2\frac{\partial f_A}{\partial \bar{\delta}_B}\frac{\partial f_A}{\partial \bar{\lambda}_B}\sigma_{\bar{\delta}_B,\bar{\lambda}_B} + 2\frac{\partial f_A}{\partial \bar{\delta}_C}\frac{\partial f_A}{\partial \bar{\lambda}_C}\sigma_{\bar{\delta}_C,\bar{\lambda}_C} \quad (10)$$

where the $\sigma_{\bar{\delta},\bar{\lambda}}$ terms represent the covariances between the mean isotopic signatures $\bar{\sigma}$ and $\bar{\lambda}$ for populations *A, B, C,* and *M*. This can be rewritten as:

$$\sigma_{fA}^{2'} = \sigma_{f_A}^2 + 2\frac{\partial f_A}{\partial \bar{\delta}_M}\frac{\partial f_A}{\partial \bar{\lambda}_M}\rho_{\delta_M\lambda_M}\sigma_{\bar{\delta}_M}\sigma_{\bar{\lambda}_M} + 2\frac{\partial f_A}{\partial \bar{\delta}_A}\frac{\partial f_A}{\partial \bar{\lambda}_A}\rho_{\delta_A\lambda_A}\sigma_{\bar{\delta}_A}\sigma_{\bar{\lambda}_A}$$
$$+ 2\frac{\partial f_A}{\partial \bar{\delta}_B}\frac{\partial f_A}{\partial \bar{\lambda}_B}\rho_{\delta_B\lambda_B}\sigma_{\bar{\delta}_B}\sigma_{\bar{\lambda}_B} + 2\frac{\partial f_A}{\partial \bar{\delta}_C}\frac{\partial f_A}{\partial \bar{\lambda}_C}\rho_{\delta_C\lambda_C}\sigma_{\bar{\delta}_C}\sigma_{\bar{\lambda}_C} \quad (11)$$

where the ρ values are the correlation coefficients between the two isotopic signatures (δ and λ) for each population (*M, A, B,* and *C*). Again letting N and D represent the numerator and denominator of Eq. 7, and substituting the values of the partial derivatives of f_A, Eq. 11 reduces to:

$$\sigma_{fA}^{2'} = \sigma_{f_A}^2 + \frac{1}{D^4}\left\{2\left[D(\bar{\lambda}_C - \bar{\lambda}_B)\right]\left[D(\bar{\delta}_B - \bar{\delta}_C)\right]\rho_{\delta_M\lambda_M}\sigma_{\bar{\delta}_M}\sigma_{\bar{\lambda}_M}\right.$$
$$+ 2\left[N(\bar{\lambda}_B - \bar{\lambda}_C)\right]\left[N(\bar{\delta}_C - \bar{\delta}_B)\right]\rho_{\delta_A\lambda_A}\sigma_{\bar{\delta}_A}\sigma_{\bar{\lambda}_A}$$
$$+ 2\left[D(\bar{\lambda}_M - \bar{\lambda}_C) - N(\bar{\lambda}_A - \bar{\lambda}_C)\right]\left[D(\bar{\delta}_C - \bar{\delta}_M)\right.$$
$$\left. - N(\bar{\delta}_C - \bar{\delta}_A)\right]\rho_{\delta_B\lambda_B}\sigma_{\bar{\delta}_B}\sigma_{\bar{\lambda}_B}$$
$$+ 2\left[D(\bar{\lambda}_B - \bar{\lambda}_M) - N(\bar{\lambda}_B - \bar{\lambda}_A)\right]\left[D(\bar{\delta}_M - \bar{\delta}_B)\right.$$
$$\left.\left. - N(\bar{\delta}_A - \bar{\delta}_B)\right]\rho_{\delta_C\lambda_C}\sigma_{\bar{\delta}_C}\sigma_{\bar{\lambda}_C}\right\} \quad (12)$$

If the correlations are assumed to be zero, these covariance terms drop out and $\sigma_{fA}^{2'}=\sigma_{f_A}^2$. As in the two-source mixing problem, approximate 95% confidence intervals for f_A can be constructed as $f_A \pm t_{.05\gamma}\sigma_{f_A}'$, and variances and confidence intervals for f_B and f_C can be calculated by switching subscripts in Eq. 9 and Eq. 12. An Excel spreadsheet to perform the calculations for the source proportions and their variances, SEs, and 95% confidence intervals for the two-source and three-source mixing models can be accessed at http://www.epa.gov/wed/pages/models.htm.

Relationship to similar mixing models

Several alternative three-source mixing models based on Euclidean distance have been used in recent studies, but these have been demonstrated to have a number of computational problems compared to the linear mixing model outlined above (Ben-David and Schell 2001; Phillips 2001) and are not recommended for use. Some authors have used the linear mixing models outlined above, but rather than algebraically determining exact solutions (Eqs. 2, 7), have used iterative procedures to find solutions with minimum residuals for either the two-source model (Brunel et al. 1995; Dawson 1998) or the three-source model (Cramer et al. 1999). If the optimization procedure works well this method should converge on the algebraic solution (zero residual) or a close approximation to it, and thus is functionally equivalent to the algebraic method. Brunel et al. (1995) also derived separate estimates of proportions for two sources from oxygen and hydrogen isotopic signatures and used a minimum residual method to find optimal compromises between the two sets of estimates.

A number of studies have used $\delta^{18}O$ and $\delta^{2}H$ to partition plant use of water from various depths. While water samples may be taken at numerous depths to determine the isotopic profile (e.g., Cramer et al. 1999), the samples must be combined into three depth classes, or two depth classes if only a single isotope is used, in order to partition water use from the different soil compartments (Walker and Richardson 1991; Cramer et al. 1999).

Sensitivity analyses

The next step was to determine the sensitivity of σ_{f_A} to a variety of factors. Clearly from Eqs. 4 and 9, σ_{f_A} depends on the variances of the mean isotopic signatures for the individual sources and the

mixture, along with the signature difference between sources, and the value of f_A. In turn, the observed variability of the mean isotopic signature for each source and mixture depends on both sampling error and analytical error, which can be examined separately in sensitivity analyses. Specifically, the variance for mean isotopic signatures for source A can be written as:

$$\sigma_{\bar{\delta}_A}^2 = \frac{\sigma_{\delta_A}^2}{n_A} + \frac{\sigma_{analytical}^2}{n_A} \tag{13}$$

where $\sigma_{\delta_A}^2$ is the population variance in isotopic signatures among individual samples of source A, $\sigma_{analytical}^2$ is the analytical error variance (i.e., variance among duplicate samples), and n_A is the number of samples taken from source A. Variances for other sources and for the mixture can be computed in the same way.

We performed sensitivity analyses on σ_{f_A} for the single isotope, two-source mixing problem by varying the levels of the following parameters: (1) the difference in isotopic signatures between the sources ($\delta_A - \delta_B$); (2) the population SDs (σ_{δ_A}, σ_{δ_B}, and σ_{δ_M}); (3) the number of samples (n_A, n_B, and n_M); (4) the analytical SD ($\sigma_{analytical}$); and (5) the value of f_A. SDs were used in (2) and (4) in order to separate out the effects of population variability and sample size, which would be confounded if SEs were used instead. Each parameter was varied individually across a range of realistic values while the other parameters were set to default values (Table 1). The population SDs (σ_{δ_A}, σ_{δ_B}, and σ_{δ_M}) were always set equal to each other and were varied together, as were the sample sizes (n_A, n_B, and n_M).

Results and discussion

Example variance and confidence interval calculations

Table 2 shows the results of applying Eq. 4 to estimate the variance of the proportion estimates for a single isotope, two-source mixing problem. The example is from Vitorello et al. (1989), who estimated the proportional contributions of the current vegetation (C_4 sugar cane) and soil from the previous vegetation (C_3 forest species) to the soil organic carbon of young (12 year) and old (50 year) tropical sugar cane fields. For the 50-year-old cane field soil, the SE for the proportion estimates calculated from Eq. 4 (2.82%), that incorporates variability in the source isotopic signatures as well as the mixture, was nearly identical to the reported SE of 2.79% (Vitorello et al. 1989) based only on the variability of the individual

Table 1 Values of parameters used in the sensitivity analyses for σ_{f_A}. Each parameter was varied individually over the ranges indicated, with the other parameters set at their default values, except that all sample sizes were run for each parameter set. Population SDs and sample sizes were the same for sources A and B, and the mixture M. The results of the sensitivity analyses are shown in Fig. 2

Parameter	Symbol	Minimum	Maximum	Default
Source difference (‰)	$\bar{\delta}_A - \bar{\delta}_B$	2	20	10
Population SD (‰)	$\sigma_{\delta_A}, \sigma_{\delta_B}, \sigma_{\delta_M}$	0.0	2.0	0.5
Sample sizes	n_A, n_B, n_M	1	10	
Analytical SD (‰)	$\sigma_{analytical}$	0.0	0.5	0.2
Source proportion	f_A	0.0	1.0	0.5

Table 2 Example of calculation of SEs for mean proportions for two sources (soil under previous C_3 forest vegetation and C_4 sugar cane as inputs into sugar cane field soil organic carbon) using a single isotopic signature ($\delta^{13}C$). Values in *regular type* are from Vitorello et al. (1989); *italicized values* were not specified in Vitorello et al. (1989) but were assigned to be consistent with oth- er values provided; and values in *bold type* are calculated from Eq. 4. The reported source proportion SE values from Vitorello et al. (1989) were determined from proportion calculations on replicate samples of the cane soil (SD from their Table 2 divided by the square root of the sample size)

	12 years cultivation			50 years cultivation		
	Cane soil	Forest soil	Sugar cane	Cane soil	Forest soil	Sugar cane
$\delta^{13}C$ [‰] (SE, SD)	−23.66 (0.16, 0.50)	−24.99 (0.09, 0.27)	−13 *(0.09, 0.30)*	−20.66 (0.33, 1.05)	−24.99 (0.09, 0.27)	−13 *(0.09, 0.30)*
Sample size	10	10	*10*	10	10	*10*
Source proportions [%] (SE) – reported		88.9 (1.28)	11.1 (1.28)		63.9 (2.79)	36.1 (2.79)
Source proportions [%] (SE) – calculated		**88.9 (1.46)**	**11.1 (1.46)**		**63.9 (2.82)**	**36.1 (2.82)**
95% Confidence limits (%)		**85.7–92.1**	**7.9–14.3**		**57.5–70.3**	**29.7–42.5**

Table 3 Example of calculation of SEs for mean proportions for three sources (moose, caribou, and salmon in wolves' diet in interior Alaska) using dual isotopic signatures (δ^{13}C and δ^{15}N). Values in *regular type* are from Szepanski et al. (1999) and values in *bold type* are calculated from Eq. 9. Food source isotopic signatures were corrected for fractionation during digestion and assimilation. The SE for δ^{13}C for caribou was reported to one decimal place as 0.0; a value of 0.04 was assigned to represent a low, but non-zero variability. The reported source proportion SE values from Szepanski et al. (1999) were determined from proportion calculations on replicate wolves (their Table 1). The source proportions from Eq. 9 differ slightly from those in Szepanski et al. (1999) since they used a different mixing model

	Wolf	Moose	Caribou	Salmon
δ^{13}C [‰] (SE)	–19.6 (0.1)	–20.7 (0.8)	–18.8 (0.04)	–18.9 (0.3)
δ^{13}N [‰] (SE)	6.4 (0.1)	4.6 (0.4)	7.0 (0.1)	14.1 (0.2)
Sample size	50	5	41	42
Source proportions [%] (SE) – reported		35.7 (2.9)	55.2 (3.0)	9.1 (0.6)
Source proportions [%] (SE) – calculated		**41.8 (18.1)**	**52.5 (24.4)**	**5.7 (6.7)**
95% Confidence limits (%)		**0–92.1**	**0–100**	**0–22.1**

mixture samples. In this case, the SEs were similar because the variation in the mixture signature (SE of 0.33‰) was 4 times greater than the variation in the source signatures (SEs of 0.09‰). For the 12-year-old cane field soil, the SE calculated from Eq. 4 (1.46%) was 14% higher than the reported SE of 1.28% (Vitorello et al. 1989). While the mixture variation (SE of 0.16‰) was still 2 times greater than the variation in the sources (0.09‰), their variability was large enough to be reflected in the SEs for source proportions. If the SEs of the source isotopic signatures had been equal to that of the mixture (0.16‰), the proportion SEs would have been 1.77%, or 38% higher than when only the mixture variability was considered.

Data from the partitioning of food sources (moose, caribou, and salmon) in interior Alaska wolves' diet using δ^{13}C and δ^{15}N (Szepanski et al. 1999) are shown in Table 3 as an example of variance calculations for a dual isotope, three-source mixing problem assuming independence of all isotopic signatures. Szepanski et al. (1999) used an alternative mixing model that gave slightly different proportion estimates (Phillips 2001), but their reported proportion SEs are used to represent the uncertainty due solely to variability in the mixture samples (individual wolves). In this example, the SEs derived from Eq. 9 that incorporate variability in both source and mixture isotopic signatures are much greater than their reported values. This is because the variations in source signatures were up to 8 times higher for δ^{13}C and 4 times higher for δ^{15}N than those of the mixture. Logically, increased uncertainty about the correct values for the source isotopic signatures should increase the uncertainty around the estimates for source proportions. However, when proportion SEs are calculated in the usual way, taking the source isotopic signatures as fixed values and computing SEs based only on variation in the mixture samples, the source variability has no effect on the calculation. Equation 4 and Eq. 9 take into account the contribution of variability both in sources and the mixture.

The situation can also be thought of as varying the location of the mixture isotopic signature within its confidence limits inside the area bounded by its sources (e.g., the wolf confidence limits inside the triangle in Fig. 1b).

When the source end-members are also allowed to move around within their confidence limits (e.g., the vertices of the triangle in Fig. 1b), the resulting variation of the source proportions is magnified. In the wolf example, the 95% confidence intervals from Eq. 9 for each of the food sources are broad and overlap zero (Table 3) as a result of this compounded uncertainty.

Sensitivity analyses

The results of our sensitivity analyses of variation in source proportion estimates in response to various factors for the single isotope, two-source case are shown in Fig. 2. For the range of values considered, the source proportion SEs were most sensitive to the difference between source signatures (Fig. 2a). Specifically, doubling the difference between sources (e.g., from 2‰ to 4‰) reduced the uncertainty of the proportion estimates by half. An example of the effect of source difference is provided by the wolf dietary study described above (Szepanski et al. 1999). In this example, the wolf and all three of its food sources were within 1.9‰ for δ^{13}C, and the wolf and two food sources are within 2.4‰ for δ^{15}N. This reduced the signal size (differentiation of the isotopic signatures of the mixture) relative to the background noise (variability in the isotopic signatures of each source) and added to the uncertainty of the proportion estimates. Similarly, because of natural variability and sampling error Hogberg (1997) recommended that quantification of N_2 fixation in ecosystem studies be attempted only when the δ^{15}N of foliage of the reference species deviates from that of N derived from N_2 fixation by >5‰. In some experimental situations, addition of isotopically labeled substrate can increase the signature difference among sources to allow better estimates of source proportions, but this problem would still exist for some study systems using isotopes at natural abundances.

The SEs of the source proportions varied linearly with the population SDs of isotopic signatures for the sources and the mixture (Fig. 2b) when the population SDs were greater than the analytical SD (0.2‰ for these analyses).

Fig. 2 Sensitivity of SEs of single isotope, two-endmember proportion estimates to **a** the difference in isotopic signatures between sources; **b** SDs for isotopic signatures in source and mixture populations; **c** analytical SD; and **d** the source proportion. Parameter values used are given in Table 1. *Curves* are shown for sample sizes of 1 (*top curve*) to 10 (*bottom curve*)
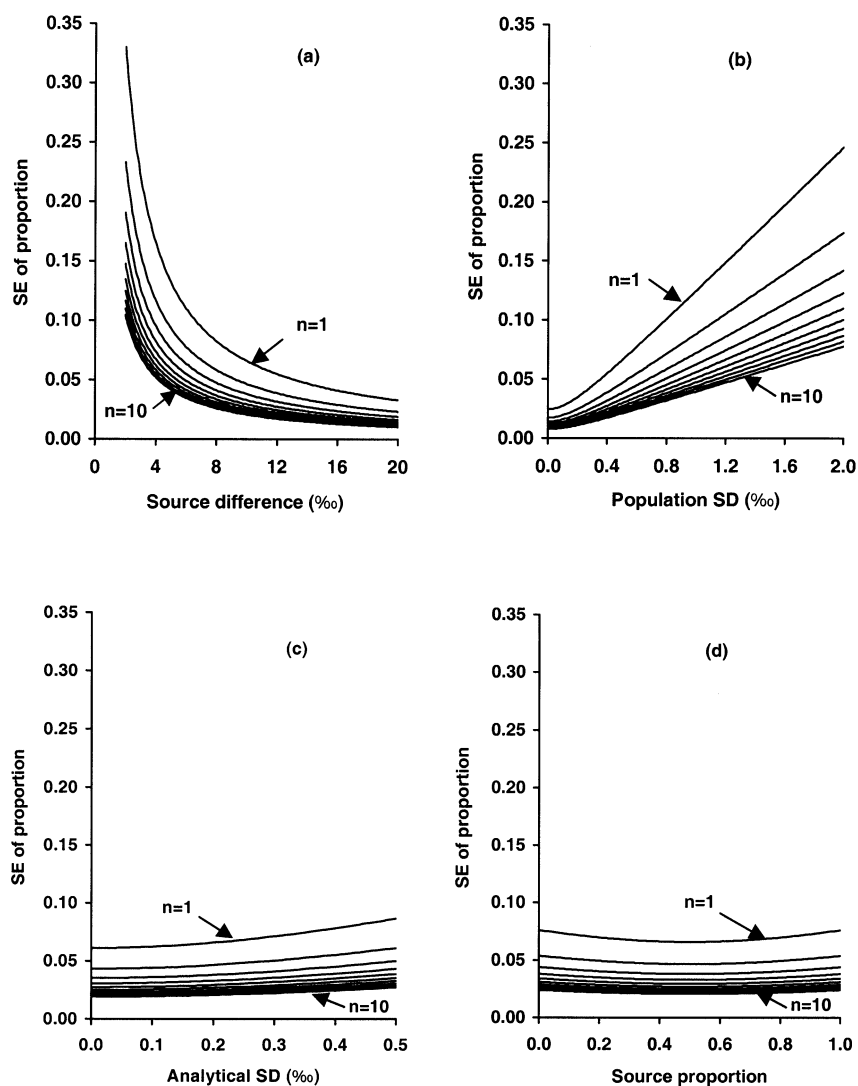
The slope of this relationship increased dramatically at smaller sample sizes. However, the variability of samples from the source and mixture populations are fixed and not under the control of the experimenter. Rather, the only way to reduce the sampling error component is to increase sample size. The variability of the proportion estimates varied inversely with the square root of sample size. For example, at a population SD of 2.0‰, the proportion SE is 0.24 for a sample size of 1, but half that size (0.12) for a sample size of 4 (Fig. 2b). This same relationship with sample size can be seen in all four graphs in Fig. 2. The study of wolf diets (Szepanski et al. 1999) also provides a good example of the importance of sample size, where the uncertainty of the mean source isotopic signatures is greatest for moose, due in large part to the 8–10 times smaller sample size than for caribou, salmon, and wolves.

The SEs of proportion estimates were not very sensitive to analytical error (Fig. 2c) over the range examined (SD of 0–0.5‰), which covers the range of typical analytical precision. This is because analytical error is gen-erally substantially smaller than the variation among, for instance, individual organisms or soil samples (Hogberg 1997). Similarly, the evenness of source proportions (e.g., 50–50 vs. 90–10) had little effect on the uncertainty of proportion estimates (Fig. 2d). Proportion SEs were smallest for proportions of 0.5, but increased only slightly as one source became dominant.

Our sensitivity analyses account for the effects of random analytical error and sampling error on proportion estimation error, but there may be additional error components which are not explicitly considered here. For example, dietary proportions require the use of isotopic signatures of the food sources that are adjusted to account for the "trophic fractionation" that occurs during digestion and assimilation. These factors are generally determined from captive feeding experiments. Some level of random error would be associated with the estimates of the fractionation factors on individual animals. There may be systematic biases as well, if the food source composition, and its consumption, digestion, and assimilation in captive feeding experiments do not match

Fig. 3 Isolines for half-width of 95% confidence intervals for two-endmember source proportions (%) as a function of single isotopic signature difference between sources (‰) and observed SDs of the isotopic signatures of the mixture and two-source populations (‰), for sample sizes of **a** *n*=3, **b** *n*=5, **c** *n*=7, and **d** *n*=10. The source proportions were fixed at 50% and the analytical SD was fixed at 0.2‰. Sample sizes were assumed to be the same for the mixture and each source. Confidence interval half-widths were calculated as $t_{.05}\sigma_{fA}$ as described in the text (see Eqs. 4, 5). In this figure, the population SDs are considered to be observed (measured) values which incorporate both natural variability and analytical error
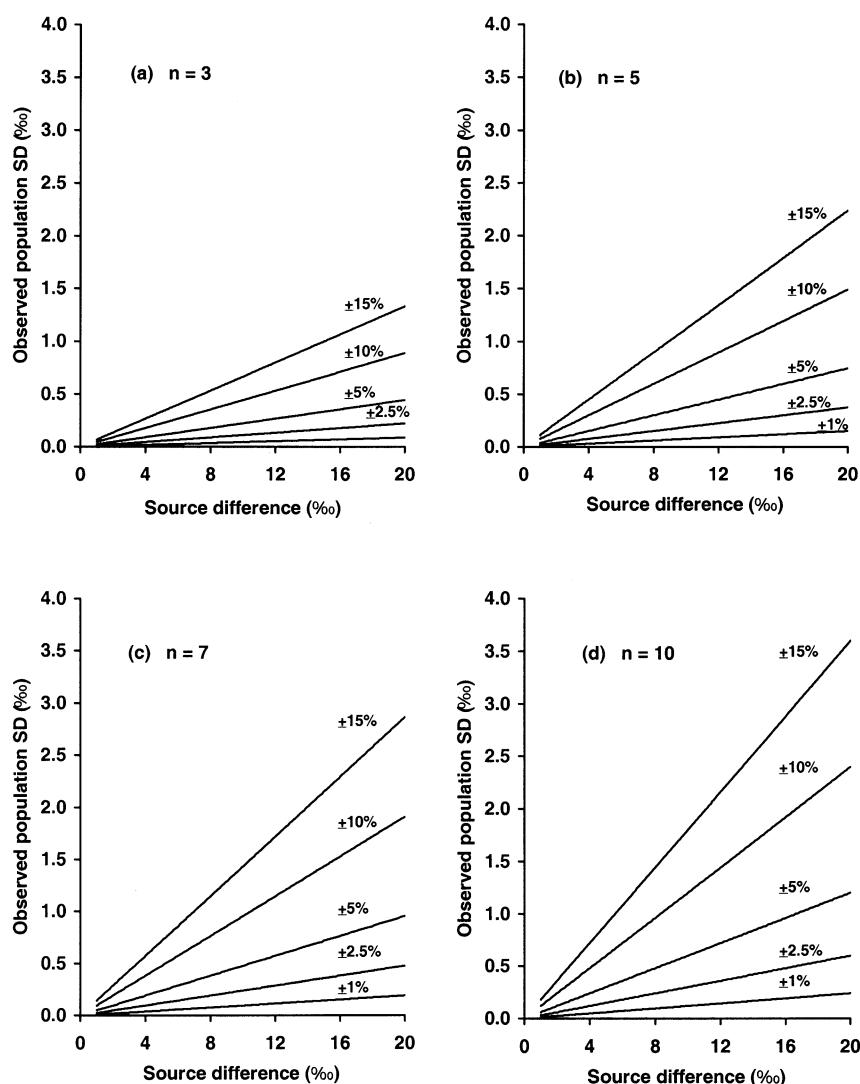
those in the wild, or if the food sources consumed represent differing proportions of the total C and N intake, for example. Other additional error components might include: systematic bias due to non-random, non-representative sampling; different animal tissues reflecting integration of dietary isotopic composition over different lengths of time; or temporal trends in source air signatures affecting changes in soil carbon.

Applications

Because of natural variability of isotopic signatures and sampling error, it has often been recommended that mixing models will work best when sources are farther apart (Dawson 1993; Hogberg 1997). Yet the utility of stable isotopes for determining the relative importance of various sources is also important for deciphering source proportions that are less divergent in isotopic composition. Examples of the need for mixing models across a relatively small span of isotopic signatures include: (1) wa-

ter source determination in the tropics where signatures vary across a small range in evaporative enrichment rather than across a large range of fractionation with the temperature of precipitation (Jackson et al. 1995); (2) use of the dual isotope approach to determine the relative importance of roots, litter, and soil organic matter ($\delta^{13}C$ range=2‰) for soil CO_2 efflux from the forest floor (Lin et al. 1999); (3) partitioning food sources that are similar in signature (Szepanski et al. 1999); and (4) determination of the relative proportion of nitrogen provided via fixation (reviewed in Hogberg 1997).

The question then, is how small a difference in source isotopic signatures is it possible to decipher? Our sensitivity analyses showed that analytical error and actual proportion had relatively little effect on source proportion error. Therefore, the minimum distance between sources depends primarily on the source and mixture SDs, the sample size, and the width of the desired confidence interval. To determine the minimum distance between source signatures that could provide sufficiently precise proportion estimates for a variety of systems, we

examined source proportion confidence intervals over a wide range of values of source signature differences and source and mixture signature SDs for four different samples sizes. The SDs for the source and mixture signatures for these analyses incorporated both natural variability and analytical error, as would be the case if one had previous signature measurements from a pilot study. The results of these analyses are shown in Fig. 3.

Comparison of source differences for various population standard deviations showed that larger source and mixture SDs could be tolerated and still achieve a desired precision when differences between sources were also large (Fig. 3). However, small differences between sources could be deciphered if the source and mixture SDs were also small, especially for large sample sizes. Furthermore, for systems with small source and mixture SDs and a large difference between source signatures, Fig. 3 also provides a basis for determining whether it is possible to reduce sample size and still attain the desired precision.

The confidence limit isolines in Fig. 3 can be used to estimate the confidence limits for a particular system. For example, although it is generally considered that isotopic differences between sources of only 2‰ would not be sufficient for linear mixing model, Fig. 3d shows that for source and mixture SDs of ±0.25‰ (which approximately correspond to a range of 1‰) and a sample size of 10 it is possible the determine the source proportion ±10% with 95% confidence. (Note that the 10% is the actual half-width of the confidence interval and not a fraction of the estimated proportion. That is, if the proportion of source A were estimated as 40%, the confidence interval would be 40%±10%, not 40%±4%.) A specific example of this situation would be the determination of the relative importance of root, litter, and soil organic matter (SOM) respiration for soil CO_2 efflux (e.g., Lin et al. 1999). Differences between these sources are generally only 2‰, but the SD of each source is also ~0.25‰ (J.W. Gregg, unpublished work). In addition to determining the relative importance of roots, litter, and SOM for soil CO_2 efflux, this information would also provide an estimate of autotrophic respiration for calculations of NPP. Therefore estimates in the range of ±10% could provide useful insight to the relative importance of the different CO_2 fluxes.

Determination of whether source proportion estimates will be ±5% or ±50% will prevent allocation of research to projects that are not feasible and promote research in areas previously not considered possible. Exact confidence limits for specific systems of interest can be calculated using the variances from Eq. 4 and Eq. 9 (or Eq. 12). Calculation of source proportion errors may be particularly useful for systems with parameter values not covered by Figs. 2, 3, e.g., where the sample sizes or SDs of the source and mixture populations are different.

Conclusions

This paper provides formulas for approximate variances, SEs, and confidence intervals for estimates of source proportions in a mixture from stable isotope analyses, for both the single isotope two-source case, and the dual isotope three-source case. (An Excel spreadsheet to perform these calculations can be accessed at http://www.epa.gov/wed/pages/models.htm.) These formulas take into account the isotopic signature variability of both the sources and the mixture. Typically when SEs of proportions are presented in the literature, they reflect only the variability of the isotopic signatures in the mixture, and use source isotopic signatures as if they are fixed values. Examination of several examples from the literature indicates that reported SEs are close to those determined from the formulas outlined here when the isotopic variability of the mixture is larger than those of the sources. However, when source variability is of equal or larger magnitude to the mixture variability, ignoring this additional component leads to substantial underestimates of the uncertainty of the source proportion estimates. In addition, sensitivity analyses indicate that the uncertainty of proportion estimates is affected most strongly by the isotopic signature difference between the sources, followed by source and mixture population variability of isotopic signatures, and sample size. Researchers can halve the uncertainty of estimates either by doubling the difference between sources (e.g., by increasing them from 2‰ to 4‰) or by quadrupling sample size (e.g., from 1 to 4). Proportion SEs were relatively unaffected by analytical error or the evenness of source proportions. A priori estimation of the confidence interval widths for source proportion estimates may expose situations where the estimates will not be of sufficient precision to be useful, thus avoiding wasting research effort. On the other hand, it may also demonstrate that reasonably precise proportion estimates are possible even when source differences are low, e.g. 2‰, if the source and mixture populations have low variability and sample size is sufficiently high. This may open up research possibilities which would have otherwise not been explored because of small differences in source signatures.

Appendix

For the two-endmember linear mixing model (Eq. 2), the solution (Eq. 4) to the general Taylor series equation for variance of f_A (Eq. 3) can be determined as follows

$$f_A = \frac{\bar{\delta}_M - \bar{\delta}_B}{\bar{\delta}_A - \bar{\delta}_B} \tag{A1}$$

The partial derivatives for f_A are:

$$\frac{\partial f_A}{\partial \bar{\delta}_M} = \frac{1}{\bar{\delta}_A - \bar{\delta}_B}$$

$$\frac{\partial f_A}{\partial \bar{\delta}_A} = \frac{(\bar{\delta}_M - \bar{\delta}_B)(-1)}{(\bar{\delta}_A - \bar{\delta}_B)^2} = \frac{-f_A}{\bar{\delta}_A - \bar{\delta}_B}$$

$$\frac{\partial f_A}{\partial \bar{\delta}_B} = \frac{(\bar{\delta}_A - \bar{\delta}_B)(-1) - (\bar{\delta}_M - \bar{\delta}_B)(-1)}{(\bar{\delta}_A - \bar{\delta}_B)^2} = \frac{\bar{\delta}_M - \bar{\delta}_B - (\bar{\delta}_A - \bar{\delta}_B)}{(\bar{\delta}_A - \bar{\delta}_B)^2}$$

$$= \frac{f_A - 1}{\bar{\delta}_A - \bar{\delta}_B} \tag{A2}$$

These derivatives are substituted into:

$$\sigma_{f_A}^2 = \left(\frac{\partial f_A}{\partial \bar{\delta}_M}\right)^2 \sigma_{\bar{\delta}_M}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_A}\right)^2 \sigma_{\bar{\delta}_A}^2 + \left(\frac{\partial f_A}{\partial \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_B}^2 \tag{A3}$$

to give:

$$\sigma_{f_A}^2 = \left(\frac{1}{\bar{\delta}_A - \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_M}^2 + \left(\frac{-f_A}{\bar{\delta}_A - \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_A}^2 + \left(\frac{f_A - 1}{\bar{\delta}_A - \bar{\delta}_B}\right)^2 \sigma_{\bar{\delta}_B}^2 \tag{A4}$$

Factoring out the common denominator this reduces to:

$$\sigma_{f_A}^2 = \frac{1}{(\bar{\delta}_A - \bar{\delta}_B)^2} \left[\sigma_{\bar{\delta}_M}^2 + f_A^2 \sigma_{\bar{\delta}_A}^2 + (1 - f_A)^2 \sigma_{\bar{\delta}_B}^2 \right] \tag{A5}$$

A similar process can be followed to derive Eq. 9 from Eqs. 7, 8 for the dual isotope, three-source linear mixing model.

References

Balesdent J, Mariotti A (1996) Measurement of soil organic matter turnover using δ^{13}C natural abundance. In: Boutton TW, Yamasaki S (eds) Mass spectrometry of soils. Marcel Dekker, New York, pp 83–111

Ben-David M, Schell DM (2001) Mixing models in analyses of diet using multiple stable isotopes: a response. Oecologia (in press). DOI 10.1007/s004420000570

Brunel J-P, Walker GR, Kennett-Smith AK (1995) Field validation of isotopic procedures for determining sources of water used by plants in a semi-arid environment. J Hydrol 167: 351–368

Cramer VA, Thorburn PJ, Fraser GW (1999) Transpiration and groundwater uptake from farm forest plots of *Casuarina glauca* and *Eucalyptus camaldulensis* in saline areas of southeast Queensland, Australia. Agric Water Manage 39:187–204

Dawson TE (1993) Water sources of plants as determined from xylem-water isotopic composition: perspectives on plant competition, distribution, and water relations. In: Ehleringer JR, Hall AE, Farquhar GD (eds) Stable isotopes and plant carbon – water relations. Academic Press, New York, pp 465–496

Dawson TE (1998) Fog in the California redwood forest: ecosystem inputs and use by plants. Oecologia 117:476–485

Griffiths H (1998) Stable isotopes, integration of biological, ecological and geochemical processes (Environmental Plant Biology series). Bios, Oxford

Hogberg P (1997) ^{15}N natural abundance in soil-plant systems (Tansley Review 95). New Phytol 137:179–203

Jackson PC, Cavelier J, Goldstein G, Meinzer FC, Holbrook NM (1995) Partitioning of water resources among plants of a lowland tropical forest. Oecologia 101:197–203

Keeling CD (1958) The concentration and isotopic abundance of atmospheric carbon dioxide in rural areas. Geochim Cosmochim Acta 13:322–334

Lin G, Ehleringer JR, Rygiewicz PT, Johnson MG, Tingey DT (1999) Elevated CO_2 and temperature impacts on different components of soil CO_2 efflux in Douglas-fir terracosms. Global Change Biol 5:157–168

Phillips DL (2001) Mixing models in analyses of diet using multiple stable isotopes: a critique. Oecologia (in press). DOI 10.1007/s004420000571

Satterthwaite FE (1946) An approximate distribution of estimates of variance components. Biom Bull 2:110–114

Szepanski MM, Ben-David M, Van Ballenberghe V (1999) Assessment of anadromous salmon resources in the diet of the Alexander Archipelago wolf using stable isotope analysis. Oecologia 120:327–335

Taylor JR (1982) An introduction to error analysis. Oxford University Press, Oxford

Veldkamp E, Weitz AM (1994) Uncertainty analysis of δ^{13}C method in soil organic matter studies. Soil Biol Biochem 26:153–160

Vitorello VA, Cerri CC, Andreux F, Feller C, Victoria RL (1989) Organic matter and natural carbon-13 distribution in forested and cultivated oxisols. Soil Sci Soc Am J 53:773–778

Walker CD, Richardson SB (1991) The use of stable isotopes in characterizing the source of water in vegetation. Chem Geol 94:145–158